

# Triadic motifs in the dependence networks of virtual societies

Wen-Jie Xie<sup>1,2</sup>, Ming-Xia Li<sup>1,2</sup>, Zhi-Qiang Jiang<sup>1,3</sup> & Wei-Xing Zhou<sup>1,2,3</sup>

April 1, 2014

<sup>1</sup>Department of Mathematics, East China University of Science and Technology, Shanghai 200237, China, <sup>2</sup>School of Business, East China University of Science and Technology, Shanghai 200237, China, <sup>3</sup>Research Center for Econophysics, East China University of Science and Technology, Shanghai 200237, China

Correspondence and requests for materials should be addressed to W.X.Z. (wxzhou@ecust.edu.cn).

(Received April 1, 2014; Accepted XXX; Published XXX)

**In friendship networks, individuals have different numbers of friends, and the closeness or intimacy between an individual and her friends is heterogeneous. Using a statistical filtering method to identify relationships about who depends on whom, we construct dependence networks (which are directed) from weighted friendship networks of avatars in more than two hundred virtual societies of a massively multiplayer online role-playing game (MMORPG). We investigate the evolution of triadic motifs in dependence networks. Several metrics show that the virtual societies evolved through a transient stage in the first two to three weeks and reached a relatively stable stage. We find that the unidirectional loop motif ( $M_9$ ) is underrepresented and does not appear, open motifs are also underrepresented, while other close motifs are overrepresented. We also find that, for most motifs, the overall level difference of the three avatars in the same motif is significantly lower than average, whereas the sum of ranks is only slightly larger than average. Our findings show that avatars' social status plays an important role in the formation of triadic motifs.**

The structure and social function of friendship networks formed of individuals with friendship ties, are of crucial importance in relation to our understanding of the evolution and behaviors of complex social systems. Conventionally, the construction of friendship networks are implemented by interviews or questionnaires [1, 2, 3]. Usually, the sizes of these networks are not large and the samples are biased. Modern technology provides us alternative ways for the collections of data on social relationships. For instance, the wide use of mobile phones leaves footprints of human's social activities, making it possible to record and analyze very large social networks [4, 5, 6, 7, 8, 9, 10]. Another important source comes from the virtual worlds resided in computer servers running massively multiplayer online role-playing games (MMORPGs). The availability of big data recorded from MMORPGs potentially enables us to test social and economic hypotheses and theories in large-scale virtual populations [11]. In recent years, such kind of data has been investigated from different angles of view [12, 13, 14, 15, 16]. In particular, studies on the structure and dynamic evolution of social networks in virtual worlds have unveiled intriguing results [17, 18, 19, 20].

In virtual worlds, avatars (virtual characters in MMORPGs) interact with each other through their social and economic activities. An avatar can propose to another avatar to make friends when they encounter. If the later accepts the proposal, they become friends and their names appear in their friend lists. Hence, the friendship is reciprocal. We will investigate the friend lists of all avatars in 248 virtual societies of a popular MMORPG (for more information about the data sets, see *Methods*). In this MMORPG, the closeness of two avatars that are friends is measured and recorded, which is termed as intimacy. If two friends finish social or economic tasks together, their intimacy will increase. Therefore, the friendship networks are weighted. Just like almost all other social networks, these friendship networks are very heterogeneous in the sense that they have broad degree distributions and that each avatar has diverse intimacies. Indeed, we find that a large proportion of friendship ties have zero intimacy, while many other ties have very large intimacy. It suggests that some friendship ties are more important and significant than others. It is thus necessary to remove insignificant ties from the friendship networks.

The simplest and natural way is to set a threshold for the intimacy [21]. If the intimacy of two friends is less than the threshold, the two avatars are not treated as friends. In this work, we adopt an alternative and more systematic way of filtering proposed by Serrano et al. [22]. The idea is to statistically validate the links of each avatar by identifying which of her links carry disproportionate fraction of the weights. Links that deviate significantly a preset null model are kept to form the “multiscale backbone” of the original network. A significant link between an avatar  $i$  and her friend  $j$  means that the link plays an important role among all her links and we can argue that avatar  $i$  depends on her friend  $j$ . This results in a directed link from  $i$  to  $j$ . We call the resulting network as dependence network, which is directed.

In this work, we investigate the evolution of triadic motifs in the dependence networks among virtual societies. We find that some motifs occur more than random while other motifs appear less than random. We also find that the avatars forming these triadic motifs often have similar levels. These features are related to the evolution of the virtual societies.

## Results

**Illustration of dependence networks, degree distributions and distribution of level differences.** For each network (a society on a given day), we construct its dependence network (see *Methods*). In Fig. 1(a) to (c), we illustrate the evolution of the dependence network in a virtual society on three days. On the second day, the number of edges is low and hence the number of triadic motifs is also low. It is consistent with the fact that significant and stable friendship relations are still infant and under development. Such dependence relationship increases with the evolution of the virtual society because collaborations among avatars increase as well as their intimacies. In addition, a link present in an early dependence network may disappear in the later network.

For a link  $i \rightarrow j$ , we define the difference of levels of the two avatars as follows

$$\Delta L_{ij} = L_j - L_i, \quad (1)$$

where  $L_i$  and  $L_j$  are the levels of avatars  $i$  and  $j$ . Figure 1(d) plots the distribution of  $\Delta L$  for all links, as well as nonreciprocal links and reciprocal links on day 30 for the same virtual society as in Fig. 1(a-c), while Fig. 1(e) shows the distributions for all societies. We find that the left part of each distribution can be approximated by an exponential distribution. The distribution for reciprocal links is symmetric in reference to  $\Delta L = 0$  as expected because both  $\Delta L_{ij}$  and  $\Delta L_{ji}$  are counted. The average proportion of reciprocal links increases from 10% to 23%. However, for nonreciprocal links, the right tail of the distribution is much fatter than the left tail. On average, about 85% of the nonreciprocal links run from low-level avatars to high-level avatars and this ratio increases along time from about 82% to 86%.














Figure 1(f) shows the distributions of in-degrees and out-degrees of the dependence network in Fig. 1(c). The in-degree distribution decays exponentially, while the out-degree distribution decays faster than an exponential. It is intriguing to observe that the in-degree distribution is much fatter than the out-degree distribution. This observation is rational and can be explained in a mechanical way. Usually, a powerful avatar is able to help many less powerful avatars and play a relative important role in the friendship lists of those less powerful avatars. However, it is difficult for an avatar to have relatively similar intimacies from many different friends.

**Distribution and evolution of motif occurrence frequency.** Let  $N_{i,s,t}$  denote the number of motif  $i$  in society  $s$  on day  $t$ . The relative occurrence frequency of motif  $i$  in virtual society  $s$  on day  $t$  is

$$O_{i,s,t} = \frac{N_{i,s,t}}{\sum_{j=1}^{13} N_{j,s,t}}. \quad (2)$$

Figure 2(a) shows the distribution of relative occurrence frequencies of the 13 motifs on day 30,  $O_{i,s,30}$ , for four typical virtual societies. We find that the distributions are quite similar for different societies and open motifs ( $M_1, M_2, M_3, M_4, M_7$  and  $M_8$ ) have higher occurrence frequency than close motifs. We also find that the occurrence frequency of motif 9 is  $O_{9,s,30} = 0$ . It suggests that the situation in which avatar  $a$  depends on avatar  $b$ , avatar  $b$  depends on avatar  $c$ , and avatar  $c$  depends on avatar  $a$  is unlikely to appear. These results are also observed for other virtual societies on other days, as shown in Fig. 2(b) for a typical society. It is found that the occurrence frequency experienced a transient stage in the first one or two weeks and then became relatively stable. For instance, the curve for  $M_4$  on the top of the plot decreases fast in the first few days and becomes almost horizontal, while many

Table 1: **Statistical significance of motif occurrences for a typical virtual society on day 30.** Without loss of clarity, we have dropped the subscripts  $s$  and  $t$  in the variables. The motifs can be classified into three groups: (1)  $M_9$  with zero occurrence; (2) Open motifs  $M_1, M_2, M_3, M_4, M_7$  and  $M_8$  that appear significantly less than random. (3) Close motifs  $M_5, M_6, M_{10}, M_{11}, M_{12}$  and  $M_{13}$  that appear significantly more than random.

$i$	$M_i$	$O_i$	$N_{i,\text{real}}$	$\langle N_{i,\text{rand}} \rangle$	$\sigma(N_{i,\text{rand}})$	$Z_i$	$p\text{-value}$
9		0.00000	0	0.10	0.30	-0.27	1.00
1		0.08358	477	622.60	0.50	-273.78	1.00
2		0.21850	1247	1339.41	0.00	-95.85	1.00
3		0.06746	385	571.70	0.80	-247.95	1.00
4		0.37112	2118	2278.60	0.60	-265.61	1.00
7		0.18311	1045	1201.70	0.60	-254.97	1.00
8		0.02541	145	270.90	0.30	-461.82	1.00
5		0.01542	88	0.30	0.50	171.89	0.00
6		0.01279	73	0.10	0.30	227.26	0.00
10		0.00088	5	0.10	0.30	19.23	0.00
11		0.01016	58	0.10	0.30	212.43	0.00
12		0.00631	36	0.10	0.30	131.74	0.00
13		0.00526	30	0.00	0.00	0.00	0.00

other curves exhibit an increasing trend in the first few days and then decrease slowly or become almost horizontal. We also show in Fig. 2(c) the evolution of the average occurrence frequency over all societies on day  $t$ :

$$O_{i,t} = \frac{1}{S} \sum_{s=1}^S O_{i,s,t}. \quad (3)$$

We confirm that there is no  $M_9$  in all dependence networks.

**Significance of motif occurrences.** Some motifs are more likely to appear than other motifs in networks, even if there are no social factors in the formation of networks. Hence, we determine the occurrence significance of motifs by the  $z$ -score

$$Z_{i,s,t} = \frac{N_{i,s,t} - \langle N_{i,s,t,\text{rand}} \rangle}{\sigma(N_{i,s,t,\text{rand}})}, \quad (4)$$

where  $N_{i,\text{rand}}$  is the occurrence number of  $M_i$  in a randomized network from the dependence network of society  $s$  on day  $t$ , and  $\langle N_{i,\text{rand},s,t} \rangle$  and  $\sigma(N_{i,\text{rand},s,t})$  are the mean and standard deviation of  $N_{i,\text{rand},s,t}$  over 100 randomized networks [23, 24].

The results of the dependence network in a typical virtual society on day 30 are shown in Table 1. Again, we find that the 13 motifs can be classified into three groups. The first group contains  $M_9$ , where  $N_{i=9,s,t} = 0$ . Although  $M_9$  was not detected in real networks, it appears occasionally in random networks. The second group contains all open motifs  $M_1, M_2, M_3, M_4, M_7$  and  $M_8$ , where  $N_{i,s,t} < \langle N_{i,s,t,\text{rand}} \rangle$ . The third group contains other close motifs  $M_5, M_6, M_{10}, M_{11}, M_{12}$  and  $M_{13}$ , where  $N_{i,s,t} > \langle N_{i,s,t,\text{rand}} \rangle$ . It is worthy stressing that  $M_{13}$  appears in real networks but not in random networks. We find that all the standard deviations are less than 1 and the  $z$ -scores are significantly different from zero. Therefore, when compared with random networks, open motifs are less likely to occur while close motifs (except  $M_9$ ) are more likely to appear in real dependent networks. For close motifs except  $M_{13}$ , the standard deviation is greater than the mean. In addition, there are five close motifs ( $M_5, M_6, M_{10}, M_{11}, M_{12}$ ) having  $\langle N_{i,\text{rand},s,t} \rangle = 0.10$  and  $\sigma(N_{i,\text{rand},s,t}) = 0.30$ . It can be understood due to the fact that, for each motif  $i$  ( $i = 5, 6, 10, 11, 12$ ), there are 10 random networks (out of 100) having  $N_{i,\text{rand},s,t} = 1$ .

Because the standard deviation can be very small or even zero when a subgraph is unlikely to occur in a network, the estimated  $z$ -score fluctuates a lot and may diverge. An alternative measure can be adopted to assess the occurrence significance of motifs

[23]. We calculate the abundance of  $M_i$  relative to random networks as follows

$$\Delta_{i,s,t} = \frac{N_{i,s,t} - \langle N_{i,s,t,\text{rand}} \rangle}{N_{i,s,t} + \langle N_{i,s,t,\text{rand}} \rangle + \varepsilon}, \quad (5)$$

where  $\varepsilon = 4$  is to ensure that  $|\Delta_i|$  is not too large when  $M_i$  appears very few times in both real and random networks, i.e., both  $N_{i,s,t}$  and  $\langle N_{i,s,t,\text{rand}} \rangle$  are close to zero [23]. This special case happens for  $M_9$ . We then compute the normalized abundance as follows

$$\text{SRP}_{i,s,t} = \frac{\Delta_{i,s,t}}{(\sum \Delta_{i,s,t}^2)^{1/2}}, \quad (6)$$

which is called the subgraph ratio profile (SRP) [23]. We can also calculate the average SRP over all virtual societies:

$$\text{SRP}_{i,t} = \frac{1}{S} \sum_{s=1}^S \text{SRP}_{i,s,t}. \quad (7)$$

The average procedure is meaningful because different societies exhibit similar motif profiles.

In Fig. 2(d), we show the subgraph ratio profile  $\text{SRP}_{i,s,t}$  for the four societies in Fig. 2(a) on day 30. The three groups identified in Table 1 are also observed. For the first group, we have  $\text{SRP}_{i=9,s,t} = 0$ . For the second group, we have  $\text{SRP}_{i,s,t} < 0$ . For the third group, we have  $\text{SRP}_{i,s,t} > 0$ . We further find that the absolute SRPs for the open motifs in the second group are less than that for the close motifs in the third group. In Fig. 2(e) and (f), we also show the evolution of subgraph ratio profiles  $\text{SRP}_{i,s,t}$  for the same society as in Fig. 2(b) and the average subgraph ratio profiles  $\text{SRP}_{i,t}$  over all societies. Again, we observe a transient stage and a relative stable stage.

**Correlation between motif counts.** We perform motif count analysis to study the correlation between motif counts in dependence networks, similar as in human cell-specific transcription factor regulatory networks [25]. Specifically, for a given day  $t$ , we plot  $N_{i,s,t}$  against  $N_{j,s,t}$  for different society  $s$ , where each  $s$  gives a point. Panels (a) and (b) of Fig. 3 show the correlation plots for  $t = 7$  and  $t = 30$ , respectively, where  $M_9$  is not included because  $N_{i=9,s,t} = 0$ . For each motif, the distribution of the motif count in the diagonal is unimodal with the peak close to the mean value. We find that the motif counts exhibit linear correlations. On average, the correlation coefficients  $\rho_{ij}$  between  $M_{10}$  and other motifs are relatively small, due to the fact that the average count of  $M_{10}$  is small with relatively large fluctuations.

In Fig. 3(c) and (d), we plot the evolution of correlation coefficient  $\rho_{i,j,t}$  for all the pairs of motifs. The correlation coefficient  $\rho_{i,j,t}$  decreases in the first two or three weeks and then increases to a relatively stable state. The count correlations between open motifs are relatively the strongest for most pairs, while the correlations between  $M_{10}$  and other motifs are relatively the weakest.

The characteristics of these correlations are closely related to the results in Fig. 2(a)-(c). It suggests that the local structure of dependence networks in diverse virtual societies exhibits universal evolution patterns. The relative ratio of motif counts  $N_{i,s,t}/N_{j,s,t}$  between  $M_i$  and  $M_j$  converges to certain constant along the evolution of virtual societies.

**Avatar levels in motifs.** Abnormal occurrence of motifs is often in relation to certain features of the involving individuals in complex social networks [26, 23, 27, 28, 10]. We now investigate the levels of avatars in triadic motifs. The levels of avatars are the most important trait, because levels reflect the power of attack, defence, and production of avatars.

Denote the levels of the three avatars in  $M_i$  by  $L_{i,1}$ ,  $L_{i,2}$  and  $L_{i,3}$ . The sum of avatar levels is

$$l_i = L_{i,1} + L_{i,2} + L_{i,3}. \quad (8)$$

The mean value of level sums of the  $N_{i,s,t}$  motifs in society  $s$  on day  $t$  can be calculated as follows,

$$l_{i,s,t} = \frac{1}{N_{i,s,t}} \sum_{k=1}^{N_{i,s,t}} l_i^{(k)}; \quad (9)$$

Similarly, we define the difference of avatar levels as

$$\Delta l_i = |L_{i,1} - L_{i,2}| + |L_{i,2} - L_{i,3}| + |L_{i,3} - L_{i,1}| = 2 [\max\{L_{i,1}, L_{i,2}, L_{i,3}\} - \min\{L_{i,1}, L_{i,2}, L_{i,3}\}], \quad (10)$$

and calculate its mean as follows

$$\Delta l_{i,s,t} = \frac{1}{N_{i,s,t}} \sum_{k=1}^{N_{i,s,t}} \Delta l_i^{(k)}. \quad (11)$$

To estimate if the sum and difference of avatar levels are significantly different from random networks (see *Methods*), we calculate the means of these two metrics for random networks, which are denoted by  $\langle l_{i,s,t,\text{rand}} \rangle$  for level sums and  $\langle \Delta l_{i,s,t,\text{rand}} \rangle$  for level difference. The  $z$ -scores of the two metrics can be obtained as follows

$$Z_{i,s,t}^l = \frac{l_{i,s,t} - \langle l_{i,s,t,\text{rand}} \rangle}{l_{i,s,t} + \langle l_{i,s,t,\text{rand}} \rangle} \quad (12)$$

and

$$Z_{i,s,t}^{\Delta l} = \frac{\Delta l_{i,s,t} - \langle \Delta l_{i,s,t,\text{rand}} \rangle}{\Delta l_{i,s,t} + \langle \Delta l_{i,s,t,\text{rand}} \rangle}. \quad (13)$$

The averaged  $z$ -scores over all virtual societies are computed as follows

$$Z_{i,t}^l = \frac{1}{S} \sum_{s=1}^S Z_{i,s,t}^l \quad (14)$$

and

$$Z_{i,t}^{\Delta l} = \frac{1}{S} \sum_{s=1}^S Z_{i,s,t}^{\Delta l}. \quad (15)$$

Figure 4 shows the evolution of average  $z$ -scores  $Z_{i,t}^l$  and  $Z_{i,t}^{\Delta l}$  for all motifs excluding  $M_0$ .

It is found that the  $Z_{i,t}^l$  values are positive. It indicates that the average level sum among avatars in the motifs is larger than the average level of all avatars. This observation can be explained by the fact that avatars in motifs often have relatively large intimacies and thus are more active in virtual societies. More activities may result in a relatively faster increase of levels. The  $Z_{i,t}^l$  curves have an overall increasing trend. Most curves increase relatively rapid in the first few weeks and then slow down, while some curves (say,  $M_1$  and  $M_3$ ) exhibit a continuous increase. It again indicates that, with the evolution of virtual societies, avatars in the motifs are even more active than other avatars. However, we notice that all the  $Z_{i,t}^l$  values are less than 0.15, suggesting that the average level of avatars is not a determining factor in the formation of dependence motifs.

On the contrary, all the  $Z_{i,t}^{\Delta l}$  values are negative, indicating that the average level difference among avatars in the motifs is smaller than that of all avatars. We can identify two groups of motifs with distinct behaviors. Three motifs ( $M_2$ ,  $M_4$  and  $M_7$ ) have relative small absolute  $Z_{i,t}^{\Delta l}$  values and the absolute  $z$ -scores decrease along time. These motifs are less significant than others, as shown by the SRP profiles in Fig. 2(d). Nine other motifs exhibit large absolute  $Z_{i,t}^{\Delta l}$  values and the absolute  $z$ -scores increase first and then decrease. In the early stage of the virtual societies, avatars are more enthusiastic to collaborate in their implementation of tasks. Although avatars are apt to make friends and collaborate with higher-level avatars, such kind of imbalanced teams with high level differences are not common unless that the higher-level avatars are very altruistic. Even such imbalanced teams form, the level difference among avatars will decrease because low-level avatars benefit more in such collaborations. This results in the increase of the absolute  $Z_{i,t}^{\Delta l}$  values. Several months after the generation of a virtual society, many avatars become less active or even inactive such that their levels increase relatively slowly, whereas the dependence motifs are persistent. The absolute  $Z_{i,t}^{\Delta l}$  values decrease.

## Discussion

We have studied the friendship networks of hundreds of virtual societies. In friendship networks, the distributions of degree and weight (measuring the closeness or intimacy between two friends) are heterogeneous. For a given avatar  $i$ , it is often true that some friends are more helpful than others. If the intimacy between  $i$  and her friend  $j$  is significantly larger than other intimacies between  $i$  and her other friends, we argue that  $i$  depends on  $j$ . Based on a filtering method [22], we construct dependence networks as the “multiscale backbone” of the original friendship networks. The dependence networks allow us to study the reciprocal and altruistic behaviors among avatars by ignoring insignificant ties that bear marginal social information.

We focused on the evolution of triadic motifs in dependence networks. Our findings show that the occurrences of motifs evolved to a stable state and are persistent. A very interesting result is that the level difference of avatars in motifs is an important factor in the formation of dependence networks, whereas the level sum of avatars are not. This finding is the same as the friendship networks of US students [3]. It suggests that virtual societies share certain common dynamics as our real society. Our work also highlights the scientific potential of virtual societies in the study of human behaviors [11].

## Methods

**Data description.** We use a huge database recorded from  $K = 124$  servers of a popular Massively Multiplayer Online Role-Playing Game (MMORPG) in China. In a virtual world residing in a server there are two opposing societies. Two avatars  $i$  and  $j$  can make friends, and a measure of closeness  $I_{ij}$  called intimacy is assigned to the friendship link. The intimacy  $I_{ij}$  can increase according to the collaborative activities of  $i$  and  $j$  if they belong to a same society; Otherwise,  $I_{ij}$  remains zero if  $i$  and  $j$  belong to two different societies. Hence the friendship networks of the two camps are essentially separate. Each avatar can maintain a friendship list, denoted as  $\mathcal{F}_i$  for avatar  $i$ . The friendship relation is symmetric: if  $i \in \mathcal{F}_j$ , then  $j \in \mathcal{F}_i$ .

**Construction of dependence network.** For a friendship network, we can construct a dependent network by filtering out insignificant edges based on a method proposed by Serrano *et al* [22]. We can construct a directed weighted network  $W$  whose element  $w_{ij}$  is the relative intimacy of avatar  $j$  in reference to all friends of avatar  $i$ :

$$w_{ij} = \frac{I_{ij}}{\sum_{j \in \mathcal{F}_i} I_{ij}}, \quad (16)$$

where  $w_{ij} \neq w_{ji}$ . Following Ref. [22], a directed link  $i \rightarrow j$  is significant at the level of  $\alpha$  if

$$\alpha_{ij} = 1 - (k_i - 1) \int_0^{w_{ij}} (1 - x)^{k_i - 2} dx < \alpha, \quad (17)$$

where  $k_i$  is the number of friends avatar  $i$  has. We use  $\alpha = 0.05$  and remove all the insignificant links, resulting a directed network. If link  $i \rightarrow j$  is significant, it means that  $j$  is relatively important to  $i$  in  $i$ 's friends. Speaking differently,  $i$  depends on  $j$  and the directed network can be termed as a dependence network.

**Definition and identification of triadic motifs.** We consider directed triadic motifs for the dependence networks, each of which contains three connected nodes. As shown in Fig. 5, there are 13 directed triadic motifs. These motifs represent different dependence structure among avatars at the microscopic level. For instance, motif 1 stands for the situation that one avatar depends on two other avatars, while motif 2 means that one avatar depends on another avatar which in turn depends on the third avatar. To identify motifs, we adopt a widely used method proposed in Ref. [26].

**Null model.** The specification of null models is very important in the assessment of certain statistic and the difference between the real networks and the reference randomized networks should focus on the factor under consideration [10]. Because we investigate the possible distinctions of the level sum and level difference of avatars in dependence motifs, we shuffle the levels of all avatars in dependence networks. Specifically, the topological structure of the dependence networks remains unchanged, while the levels of avatars are randomized by repeatedly exchange the levels of two arbitrarily chosen avatars.

## References

- [1] Currarini, S., Jackson, M. O. & Pin, P. An economic model of friendship: homophily, minorities and segregation. *Econometrica* **77**, 1003–1045 (2009).
- [2] Currarini, S., Jackson, M. O. & Pin, P. Identifying the roles of race-based choice and chance in high school friendship network formation. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 4857–4861 (2010).
- [3] Ball, B. & Newman, M. E. J. Friendship networks and social status. *Network Science* **1**, 16–30 (2013).
- [4] Palla, G., Barabási, A.-L. & Vicsek, T. Quantifying social group evolution. *Nature* **446**, 664–667 (2007).

- [5] Onnela, J.-P. *et al.* Structure and tie strengths in mobile communication networks. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 7332–7336 (2007).
- [6] Kumpula, J. M., Onnela, J.-P., Saramäki, J., Kaski, K. & Kertész, J. Emergence of communities in weighted networks. *Phys. Rev. Lett.* **99**, 228701 (2007).
- [7] Eagle, N., Pentland, A. & Lazer, D. Inferring friendship network structure by using mobile phone data. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 15274–15278 (2009).
- [8] Jo, H.-H., Pan, R. K. & Kaski, K. Emergence of bursts and communities in evolving weighted networks. *PLoS One* **6**, e22687 (2011).
- [9] Jiang, Z.-Q. *et al.* Calling patterns in human communication dynamics. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 1600–1605 (2013).
- [10] Kovanen, L., Kaski, K., Kertész, J. & Saramäki, J. Temporal motifs reveal homophily, gender-specific patterns, and group talk in call sequences. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 18070–18075 (2013).
- [11] Bainbridge, W. S. The scientific research potential of virtual worlds. *Science* **317**, 472–476 (2007).
- [12] Jiang, Z.-Q., Zhou, W.-X. & Tan, Q.-Z. Online-offline activities and game-playing behaviors of avatars in a massive multiplayer online role-playing game. *EPL (Europhys. Lett.)* **88**, 48007 (2009).
- [13] Jiang, Z.-Q., Ren, F., Gu, G.-F., Tan, Q.-Z. & Zhou, W.-X. Statistical properties of online avatar numbers in a massive multiplayer online role-playing game. *Physica A* **389**, 807–814 (2010).
- [14] Thurner, S., Szell, M. & Sinatra, R. Emergence of good conduct, scaling and Zipf laws in human behavioral sequences in an online world. *PLoS One* **7**, e29796 (2012).
- [15] Szell, M., Sinatra, R., Petri, G., Thurner, S. & Latora, V. Understanding mobility in a social petri dish. *Sci. Rep.* **2**, 457 (2012).
- [16] Szell, M. & Thurner, S. Social dynamics in a large-scale online game. *Adv. Complex Sys.* **15**, 1250064 (2012).
- [17] Szell, M., Lambiotte, R. & Thurner, S. Multirelational organization of large-scale social networks in an online world. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 13636–13641 (2010).
- [18] Szell, M. & Thurner, S. Measuring social dynamics in a massive multiplayer online game. *Soc. Networks* **32**, 313–329 (2010).
- [19] Klimek, P. & Thurner, S. Triadic closure dynamics drives scaling laws in social multiplex networks. *New J. Phys.* **15**, 063008 (2013).
- [20] Szell, M. & Thurner, S. How women organize social networks different from men. *Sci. Rep.* **3**, 1214 (2013).
- [21] Xie, W.-J. *et al.* Division of labor, skill complementarity, and heterophily in socioeconomic networks. submitted (2014).
- [22] Serrano, M. Á., Boguñá, M. & Vespignani, A. Extracting the multiscale backbone of complex weighted networks. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 6483–6488 (2009).
- [23] Milo, R. *et al.* Superfamilies of evolved and designed networks. *Science* **303**, 1538–1542 (2004).
- [24] Milo, R., Kashtan, N., Itzkovitz, S., Newman, M. E. J. & Alon, U. Uniform generation of random graphs with arbitrary degree sequences (2004). <http://arxiv.org/abs/cond-mat/0312028>.
- [25] Tran, N. H., Choi, K. P. & Zhang, L.-X. Counting motifs in the human interactome. *Nat. Commun.* **4**, 2241 (2013).
- [26] Milo, R. *et al.* Network motifs: Simple building blocks of complex networks. *Science* **298**, 824–827 (2002).

- [27] Kovanen, L., Karsai, M., Kaski, K., Kertész, J. & Saramäki, J. Temporal motifs in time-dependent networks. *J. Stat. Mech.* **2011**, P11005 (2011).
- [28] Jiang, Z.-Q. *et al.* Trading networks, abnormal motifs and stock manipulation. *Quant. Finance Lett.* **1**, 1–8 (2013).

## Acknowledgements

We acknowledge financial support from the National Natural Science Foundation of China (11205057 and 11375064), the Humanities and Social Sciences Fund of the Ministry of Education of China (09YJCZH040), the Ph.D. Programs Foundation of Ministry of Education of China (20120074120028), and Fundamental Research Funds for the Central Universities.



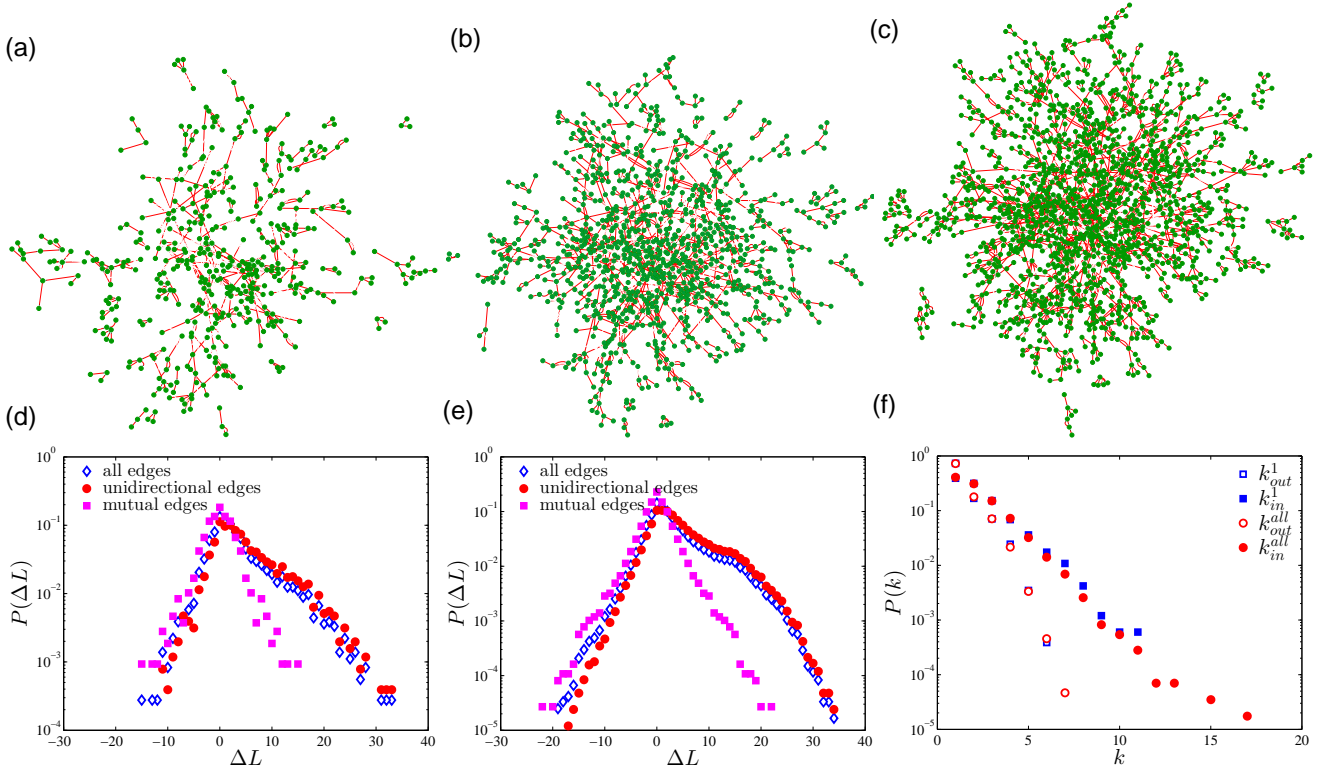
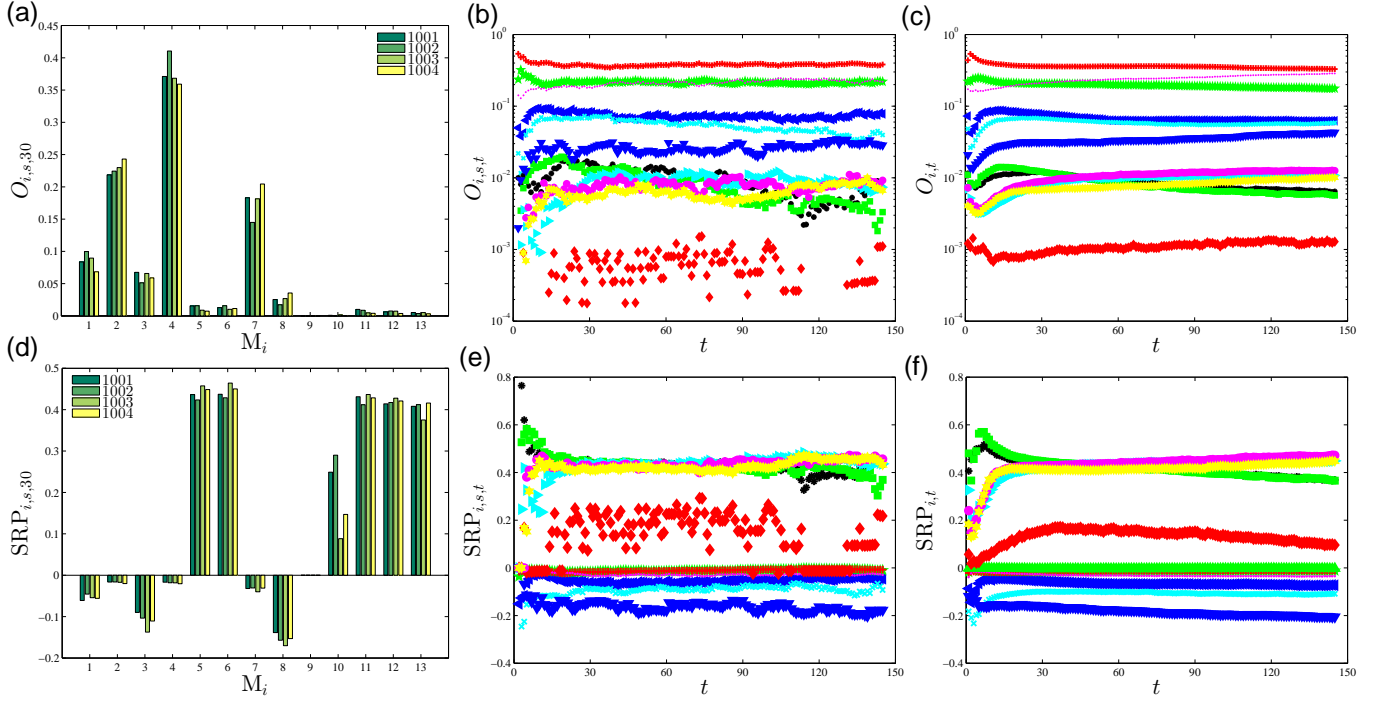
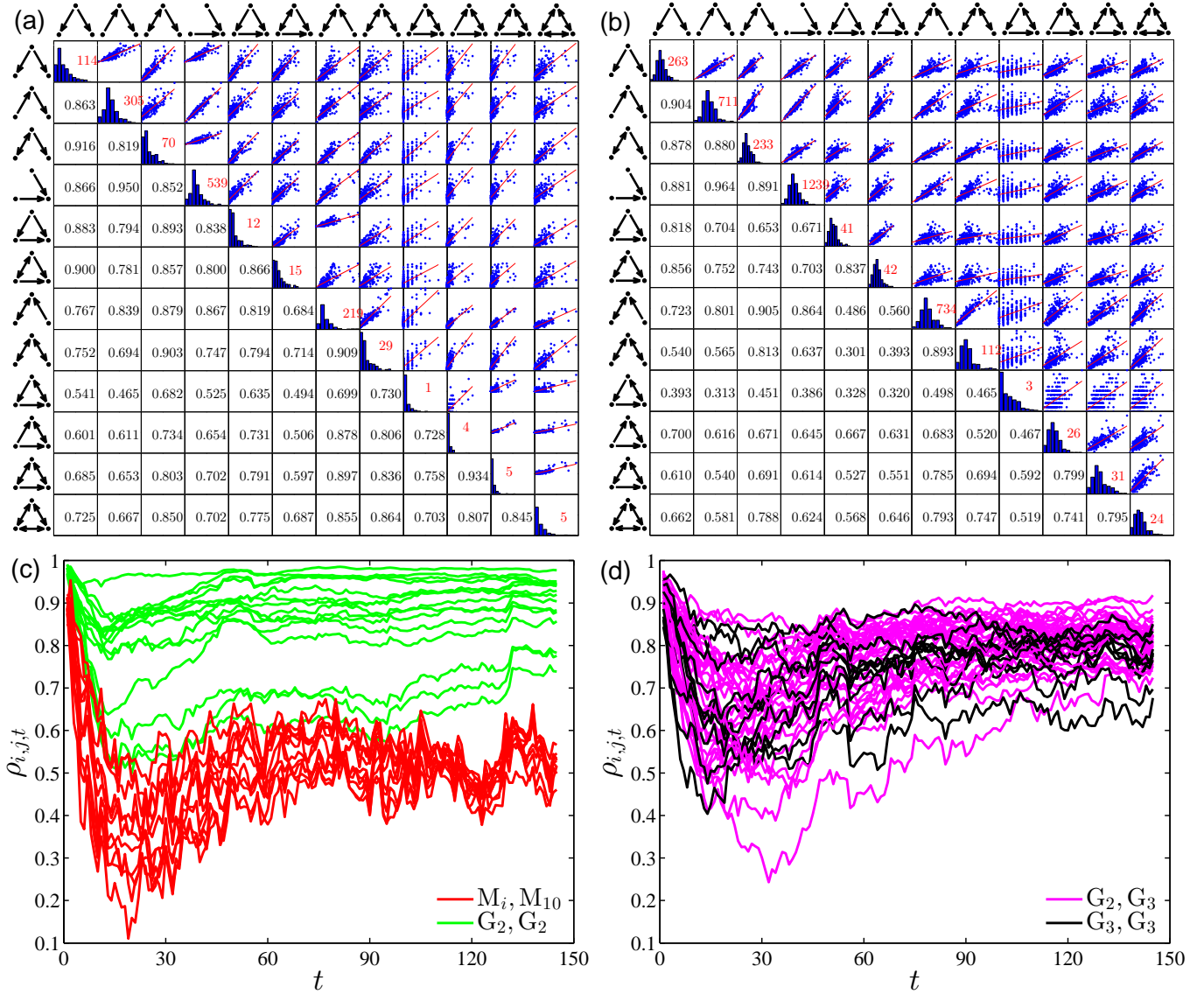


Figure 1: **Illustration of dependence networks, degree distributions and distribution of level differences.** (a) The dependence network of a virtual society on day 2. (b) The dependence network of the same virtual society on day 5. (c) The dependence network of the same virtual society on day 30. (d) Distributions of  $\Delta L_{ij}$  for all links, for nonreciprocal links (link  $i \rightarrow j$  exists while link  $j \rightarrow i$  does not exist), and for reciprocal links (both links  $i \rightarrow j$  and  $j \rightarrow i$  exist) on day 30 for the same virtual society. (e) Distributions of  $\Delta L_{ij}$  for all links, for nonreciprocal links, and for reciprocal links on day 30 for all virtual society. (f) In-degree distribution and out-degree distribution for the same virtual society and for all societies.



**Figure 2: Distribution of triadic motifs.** (a) Occurrence frequencies  $O_{i,s,30}$  of the 13 motifs on the 30th day ( $t = 30$ ) for four virtual societies. (b) Evolution of occurrence frequencies  $O_{i,s,t}$  of the 13 triadic motifs for a typical virtual society  $s$ . The occurrence frequency of motif 9 is zero along time, that is  $O_{9,s,t} = 0$ . (c) Evolution of average occurrence frequencies  $O_{i,t} = \frac{1}{S} \sum_{s=1}^S O_{i,s,t}$  of the 13 triadic motifs over all virtual society. The average occurrence frequency of motif 9 is zero along time, that is  $O_{9,t} = 0$ . (d) Significance  $SRP_{i,s,30}$  of the 13 motifs on the 30th day ( $t = 30$ ) for four virtual societies. (e) Evolution of significance  $SRP_{i,s,t}$  of the 13 triadic motifs for a typical virtual society  $s$ . (f) Evolution of average significance  $SRP_{i,t} = \frac{1}{S} \sum_{s=1}^S SRP_{i,s,t}$  of the 13 triadic motifs over all virtual society. The correspondence between the colourful symbols in plots (b), (c), (e) and (f) and motifs are given in Fig. 5.



**Figure 3: Correlation of motif counts in 248 dependence networks.** (a) The panels in the upper triangle are scatter plots of motif counts between 13 motif counts on day 7. There are no  $M_9$  motifs detected in the dependence networks. The correlation coefficients of motifs counts are presented in the lower triangular panels. Each diagonal panel shows the count distribution of these motifs and also the average motif count. (b) Same as (a) for day 30. (c) Evolution of motif count correlation coefficient  $\rho_{i,j,t}$  between  $M_{10}$  and other motifs (red curves) and between open motifs in the second group (green curves). (d) Evolution of motif count correlation coefficient  $\rho_{i,j,t}$  between close motifs in the third group (black curves) and between open motifs in the second group and close motifs in the third group (magenta curves).

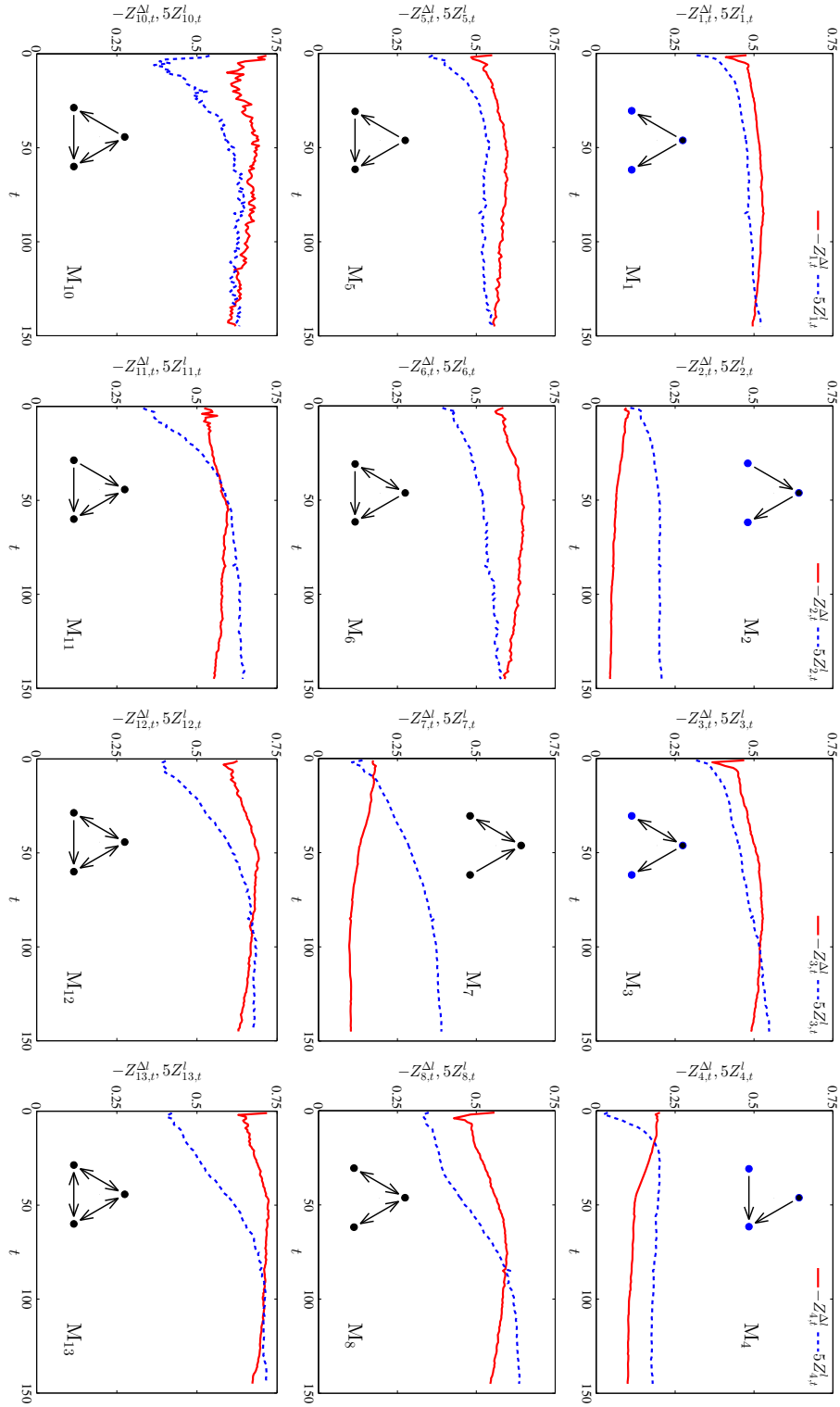


Figure 4: **Evolution of  $z$ -scores of the average level sum and level difference of avatars in motifs.** Each plot corresponds to a motif. For clarity, the  $Z_{i,t}^l$  values has been multiplied by a factor of 5.

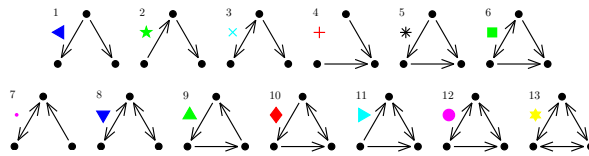


Figure 5: **List of triadic dependence motifs.** A unidirectional arrow from  $i$  to  $j$  means that avatar  $i$  depends on avatar  $j$ , while a bidirectional arrow indicates that the two avatars are mutually dependent.